Multi-Model Ensemble Wake Vortex Prediction

Stephan Körner *, Frank Holzäpfel *, Nash'at Ahmad [+]

* German Aerospace Center (DLR)
Institute of Atmospheric Physics

+ NASA Langley Research Center

WakeNet 2015, Amsterdam
22.04.2015

vortex position

decay

APA3.2

circulation

NASA

TDP2.1

D2P

DLR

APA3.4

ground effect

Multi-Model Ensemble Wake Vortex Prediction

probabilistic forecast

Bayesian Model Averaging

wake vortex behaviour

Reliability Ensemble Averaging

deterministic forecast

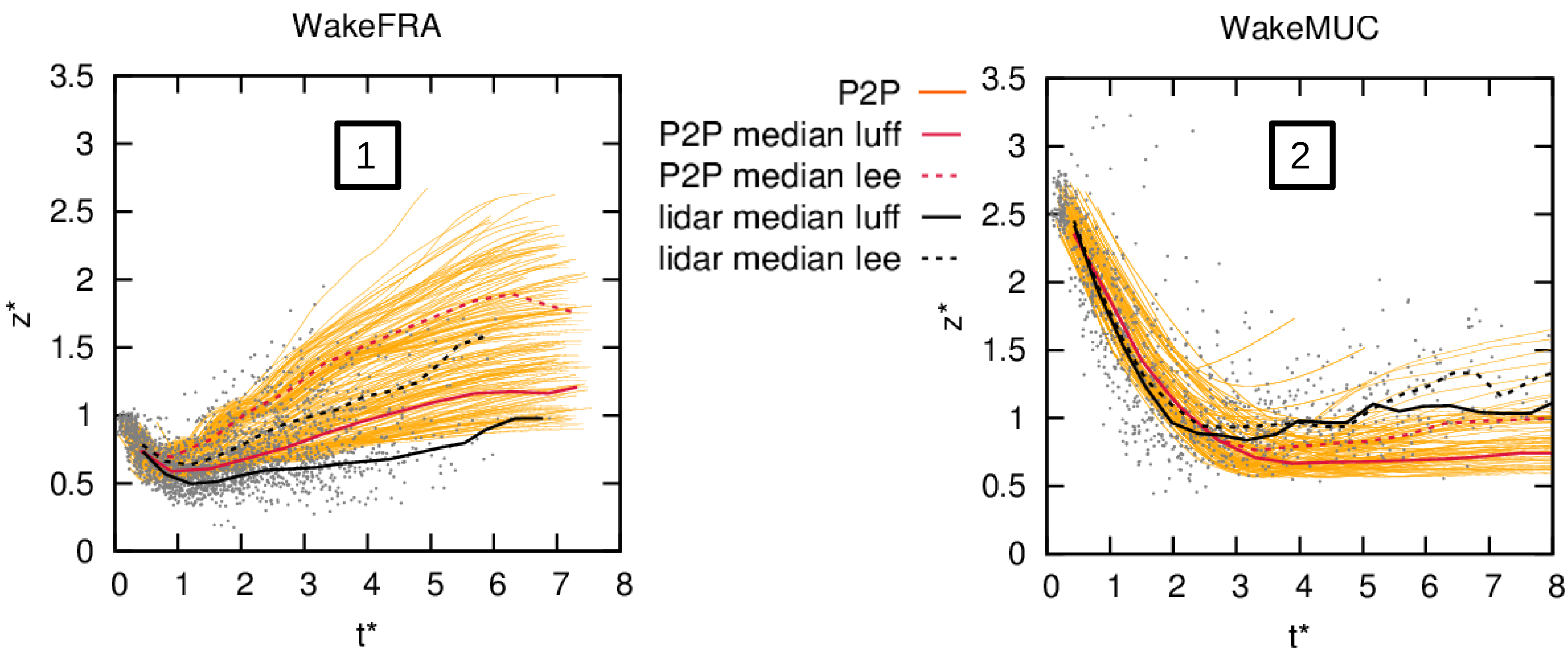
en-route

airports

outperform individual models

Revision of P2P - Motivation

wake vortex descent



88 landings

$\Gamma^*=\Gamma/\Gamma_0$, $z^*=z/b_0$, $y^*=y/b_0$, $t^*=t/t_0$

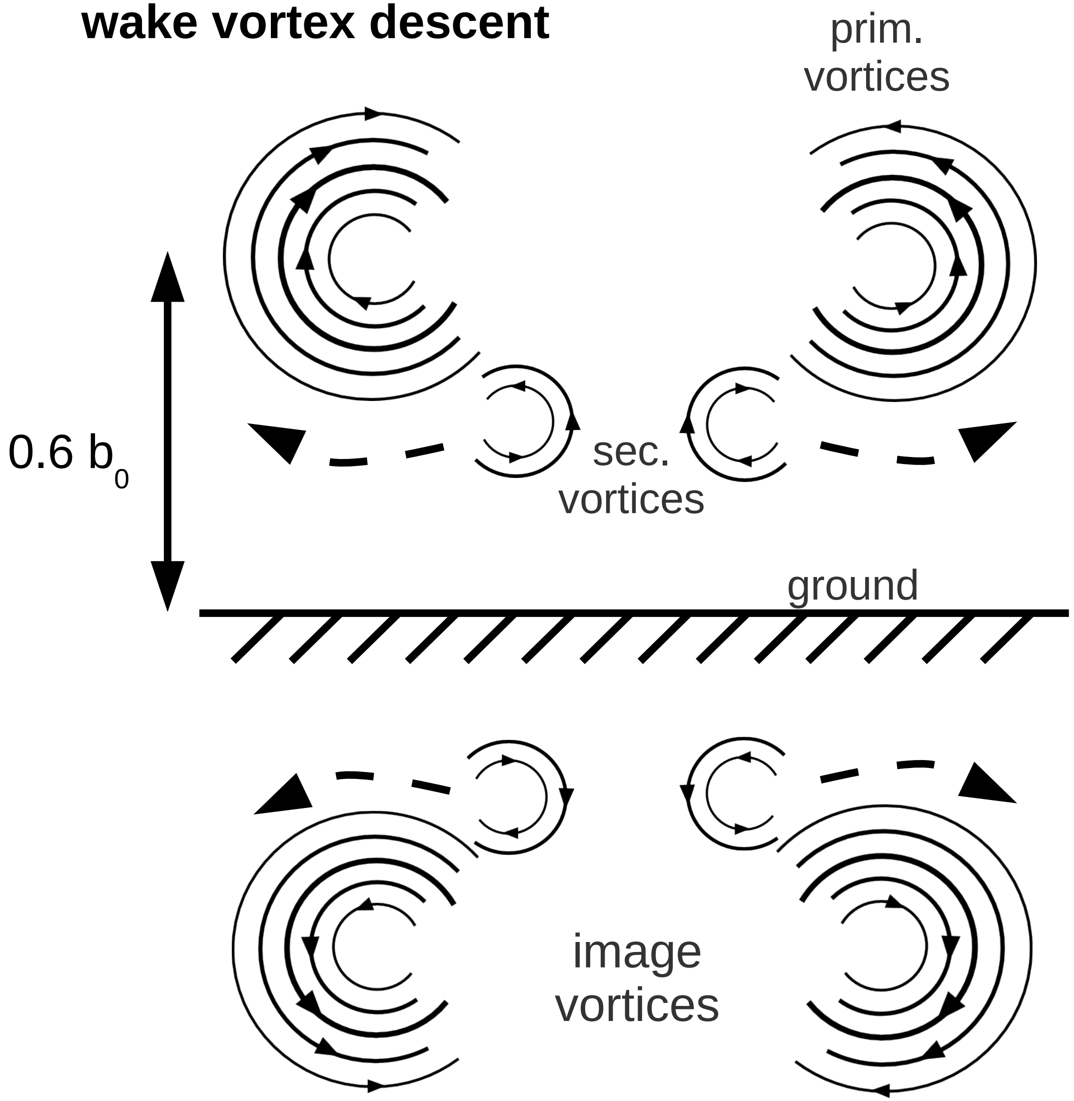
b_0 = initial vortex spacing

w_0 = initial vortex descent speed

70 landings

Revision of P2P

wake vortex descent



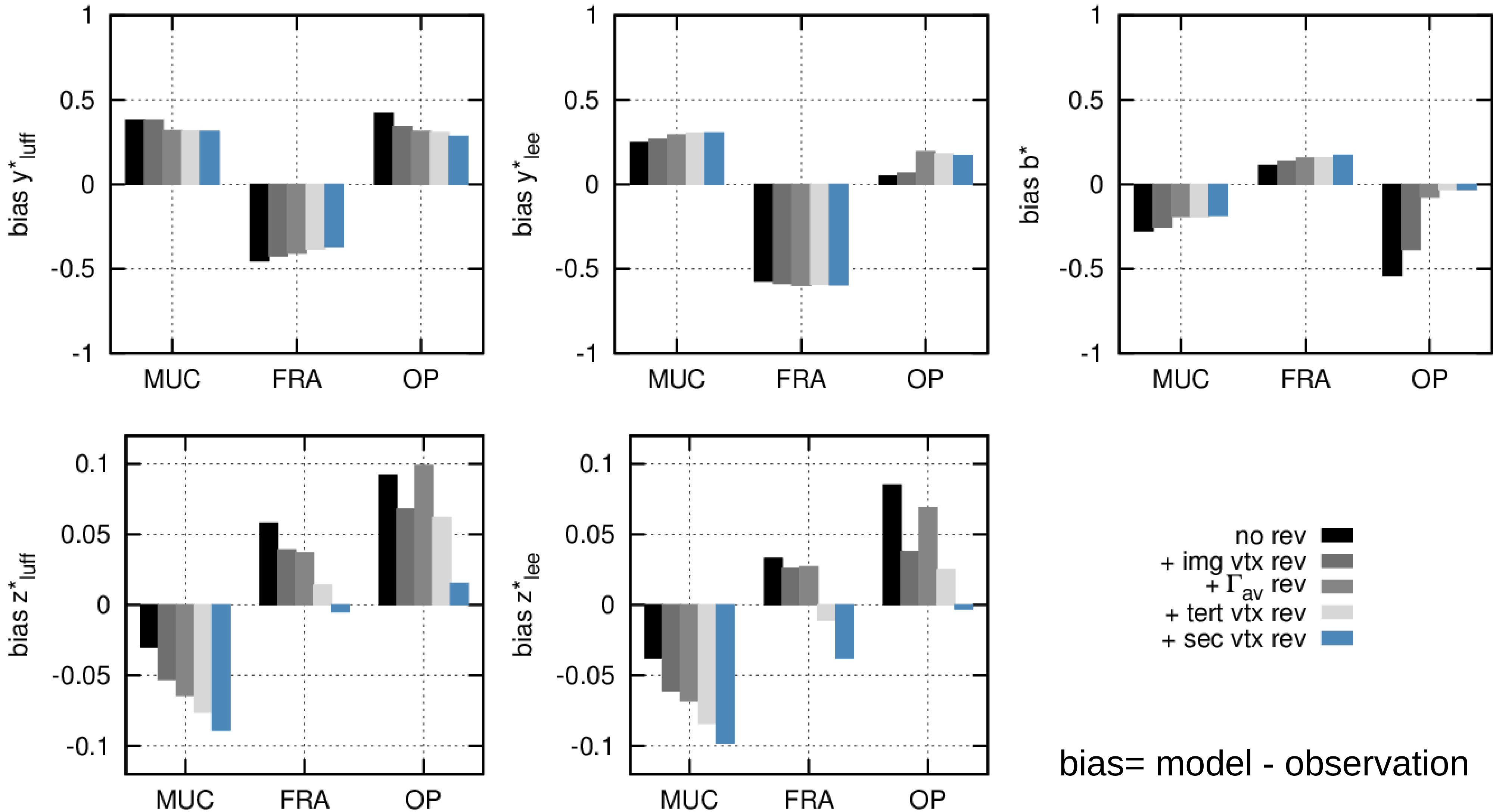
1

- secondary vortices weakened by 30 % after first round
- tertiary vortices weakened by 30 % from the beginning ($0.7 * \Gamma_{sec}$)

2

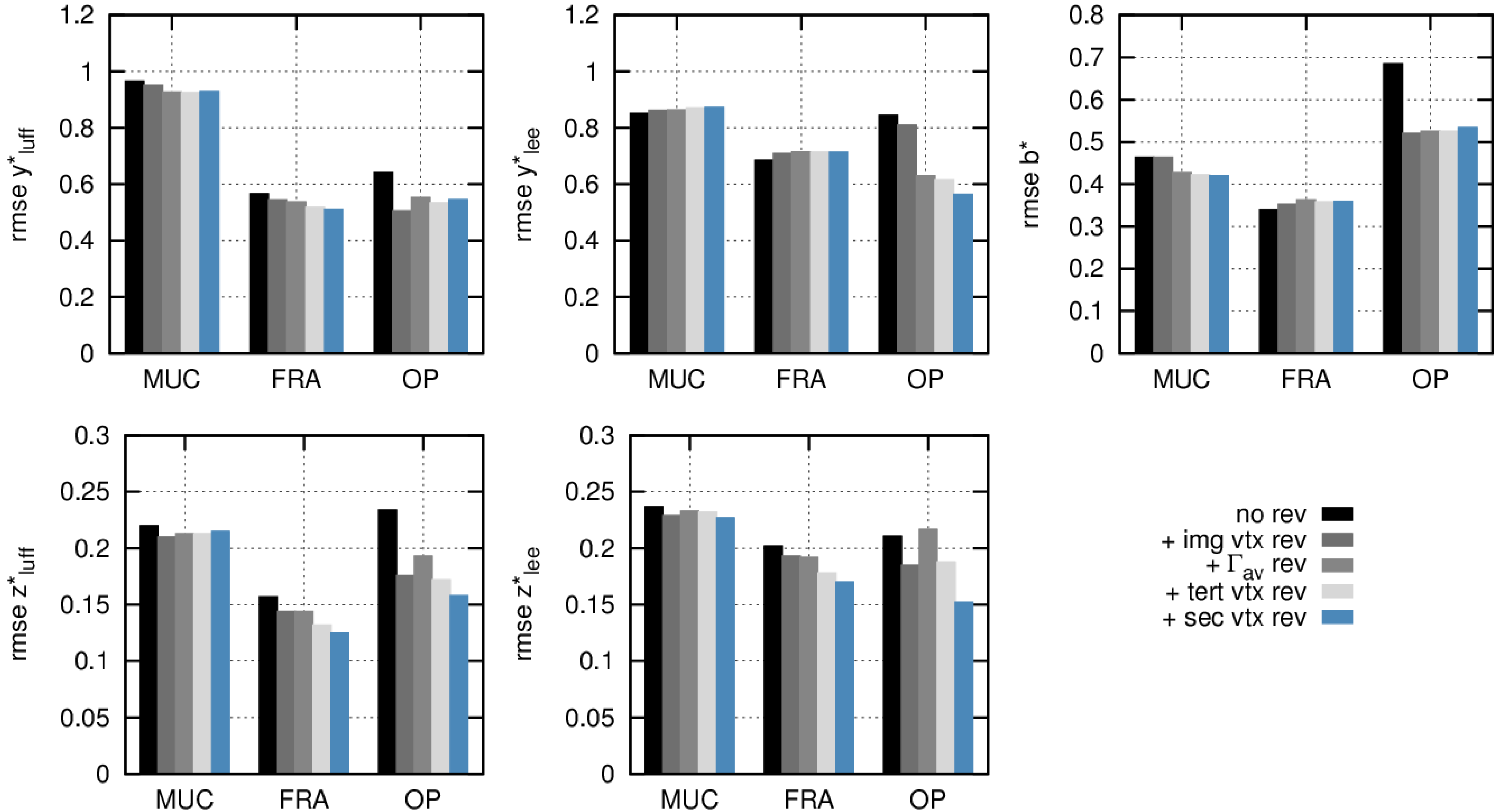
- vortex-ground interaction above 0.6 b_0: not yet further investigated
- vortex ground interaction not only distance but also time dependent ?

Revision of P2P



bias= model - observation

Revision of P2P



Multi Model Ensemble

Sugar

How to mix several good ingredients?

Water

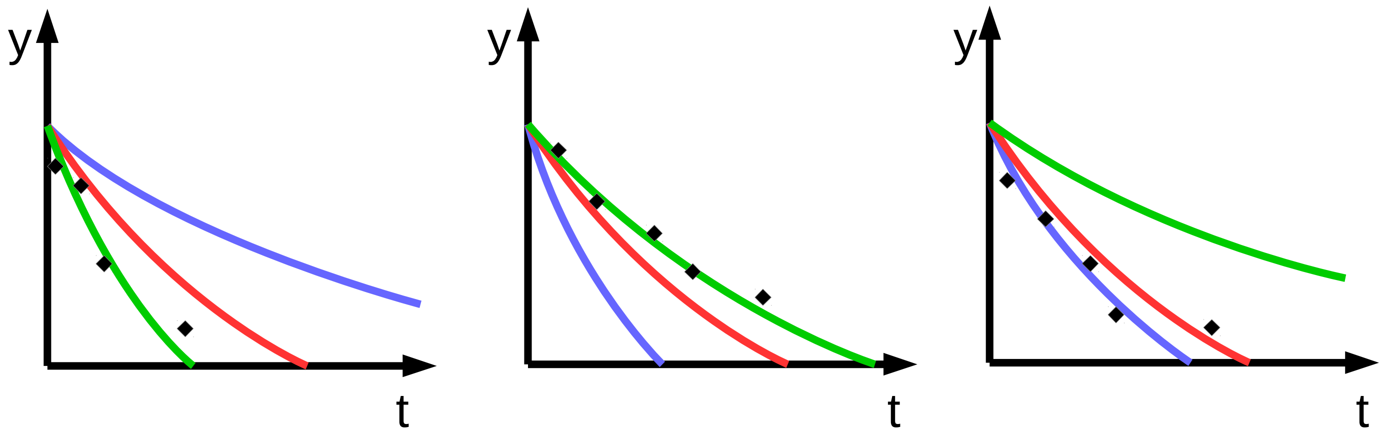
Lemon Juice

Lemonade

Why not use the best ensemble member exclusively?

Why not use the best ensemble member exclusively?

- which is the best member?
- in average best performing member can sometimes be the worst one



Can an ensemble outperform its best member?

- success of ensemble appr.: any model can be the best sometimes
- consistently low performing models → no increase of skill

Yes!

Hagedorn et al., 2005

Ensemble Members

NASA-DLR cooperation

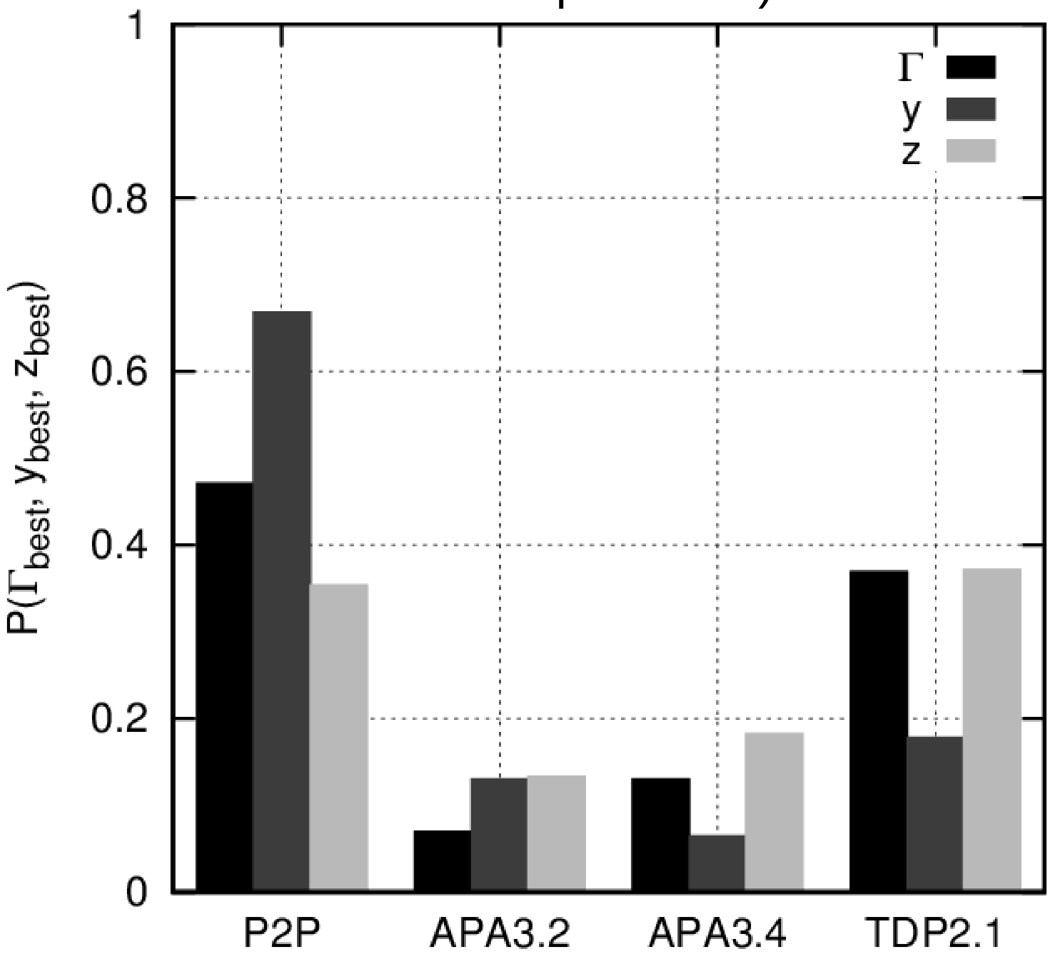
D2P
- deterministic output of P2P
- based on decaying potential vortex, adapted to LES results (DLR)

TDP 2.1
- considers effect of crosswind shear on vortex descent (NASA)

APA 3.2
- decay and transport model according to Sarpkaya (NASA)

APA 3.4
- reduced effect of stratification (NASA)

Probability that one of the models delivers the best forecast
(in ground-effect, on the basis of rmse for 99 example cases)



Multi-Model Ensemble

Reliability Ensemble Averaging (REA)

$$\tilde{f} = \tilde{A}(f) = \frac{\sum_i R_i f_i}{\sum_i R_i}$$

iteration loop

Giorgi and Mearns, 2002

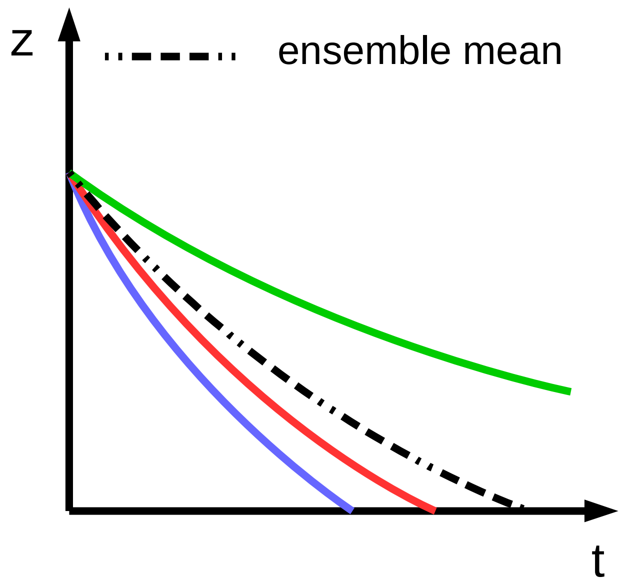
$$R_i = [(\underbrace{R_{B,i}}_{\text{model performance (a-priori)}})^m \cdot (\underbrace{R_{D,i}}_{\text{model convergence}})^n]^{[1/(m \cdot n)]} = \left\{ \left[\frac{nv}{abs(B_i)} \right]^m \left[\frac{nv}{abs(D_i)} \right]^n \right\}^{[1/(m \cdot n)]}$$

f_i	=	forecast of model i
$\tilde{A}(f)$	=	REA-forecast
R_i	=	reliability factor of model i
nv	=	natural variability
B_i	=	absolute bias of model i
D_i	=	absolute difference between forecast of model i and ensemble mean

Multi-Model Ensemble

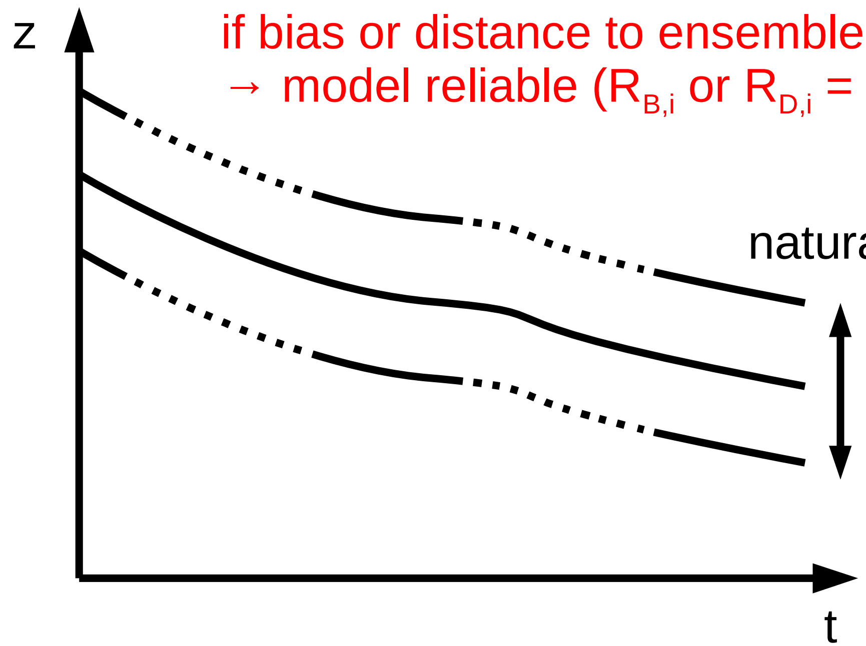
Reliability Ensemble Averaging

$R_{D,i}$ depends on distance to ensemble mean:



natural variability

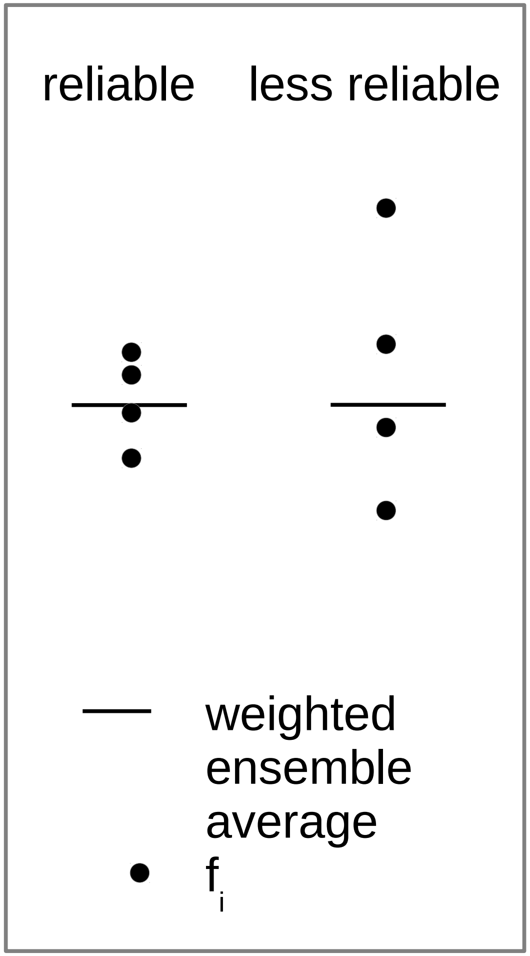
if bias or distance to ensemble mean < nv → model reliable ($R_{B,i}$ or $R_{D,i}$ = 1)



→ **nv = model resolution limit**

Multi-Model Ensemble

Reliability Ensemble Averaging uncertainty bounds:

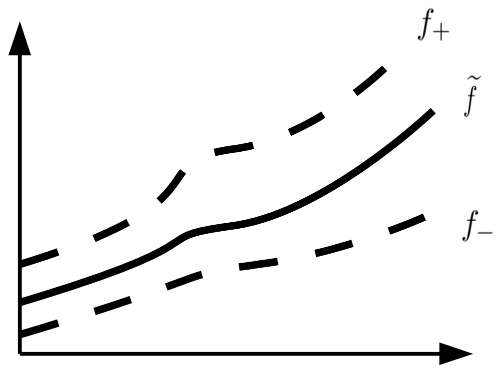


$$\tilde{\delta}_f = [\tilde{A}(f_i - \widetilde{f})^2]^{1/2} = \left[\frac{\sum_i R_i (f_i - \widetilde{f})^2}{\sum_i R_i} \right]^{1/2}$$

uncertainty bounds depend on ensemble spread

$$f_+ = \widetilde{f} + \tilde{\delta}_f$$

$$f_- = \widetilde{f} - \tilde{\delta}_f$$



according to Giorgi and Mearns, 2002

Application to Wake Vortex Models

Reliability Ensemble Averaging

Training

- mixture of landings from WakeFRA, WakeMUC and WakeOP
- 95 selected cases

$R_{B,i}$ and $R_{D,i}$

- $R_{B,z,i}(t)$, $R_{B,y,i}(t)$, $R_{B,\Gamma,i}(t)$, $R_{D,z,i}(t)$, $R_{D,y,i}(t)$, $R_{D,\Gamma,i}(t)$
- $\Delta t^*=2\ t_0$
- separately for luff and lee vortices
- weights for reliability factors: $R_{B,z,i}$: m=1.0, $R_{D,z,i}$: n=0.3

Uncertainty envelope

- initial condition uncertainty added (not considered in original approach):

variable	unit	σ (standard deviation)
true airspeed	[m/s]	4
air density	[kg/m³]	0.0048
weight	[kg]	1300
z0	[m]	7
y0	[m]	25

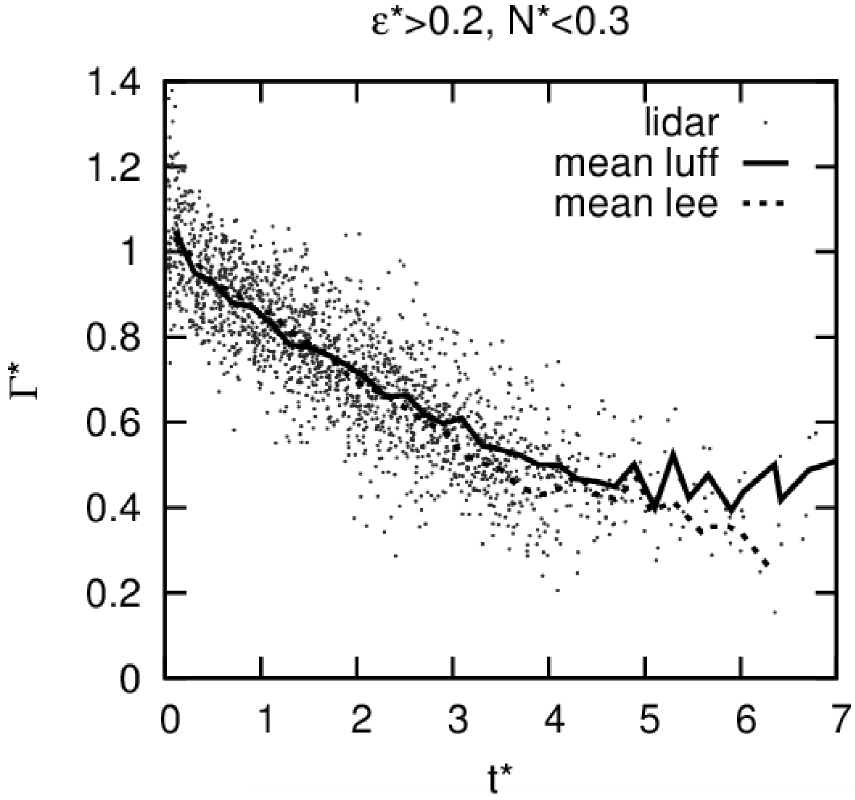
variable	unit	σ (standard deviation)
if initial conditions derived from lidar:		
z0	[m]	9
y0	[m]	13
Γ_0	[m²/s]	13

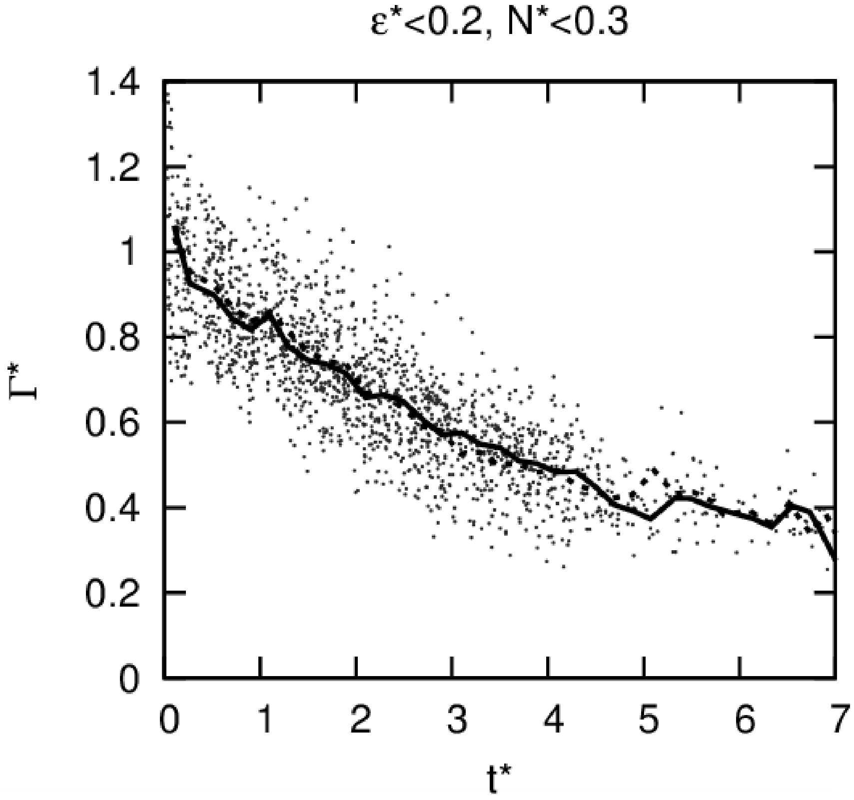
Application to Wake Vortex Models

REA natural variability, Γ*

N^*	$= N^* t_0$
ε^*	$= (\varepsilon\, {}^* b_0)^{1/3}/w_0$
v^*	$= v/w_0$

	$N^* < 0.3,\ \epsilon^* > 0.25$	$N^* < 0.3,\ \epsilon^* < 0.25$
$\sigma_{obs,\Gamma^*_{luff}}$	0.097 (1073)	0.094 (956)
$\sigma_{obs,\Gamma^*_{lee}}$	0.081	0.096



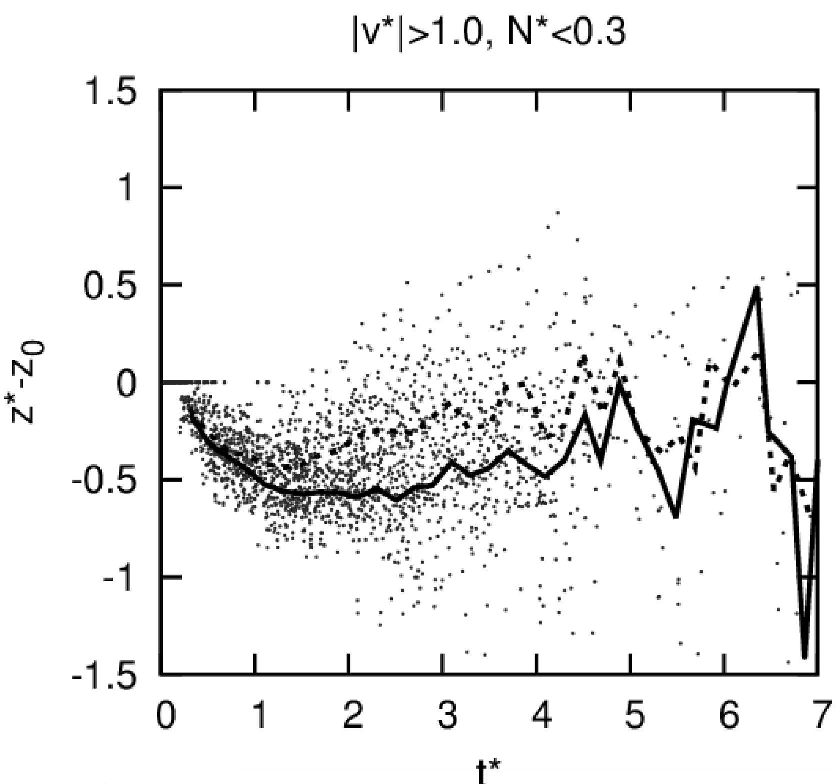


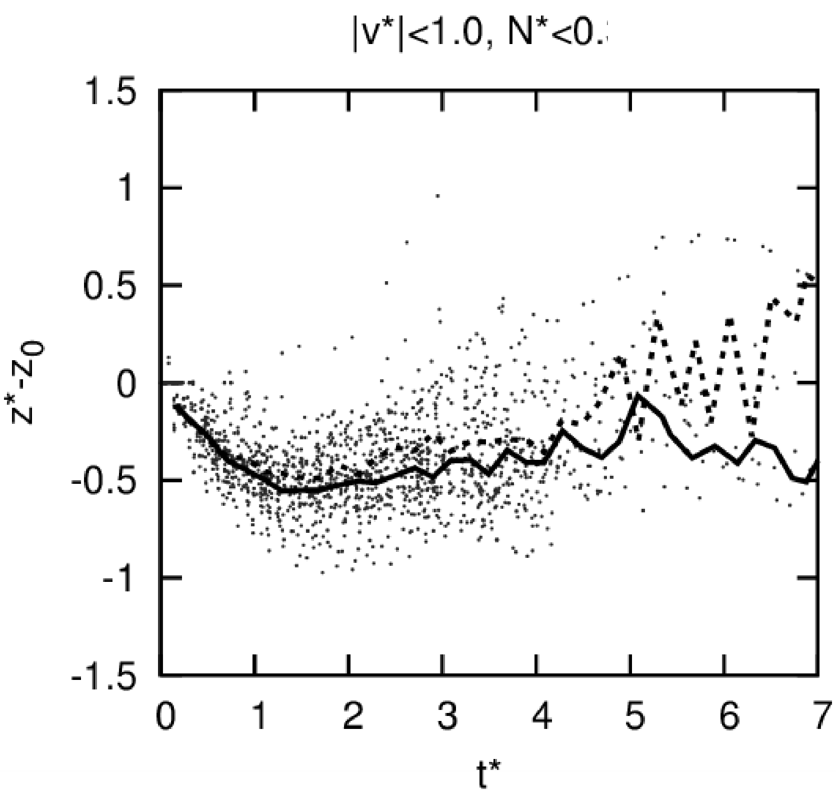
Application to Wake Vortex Models

REA natural variability, z*

$$\sigma_{obs} = \sqrt{\sigma_{err}^2 + \sigma_{nv}^2}$$

	$N^* < 0.3$, $\|v^*\| > 0.5$	$N^* < 0.3$, $\|v^*\| < 0.5$
$\sigma_{obs,z^*_{luff}}$	0.35 (1224)	0.200 (805)
$\sigma_{obs,z^*_{lee}}$	0.35	0.262





Results

REA forecast

(one single landing)

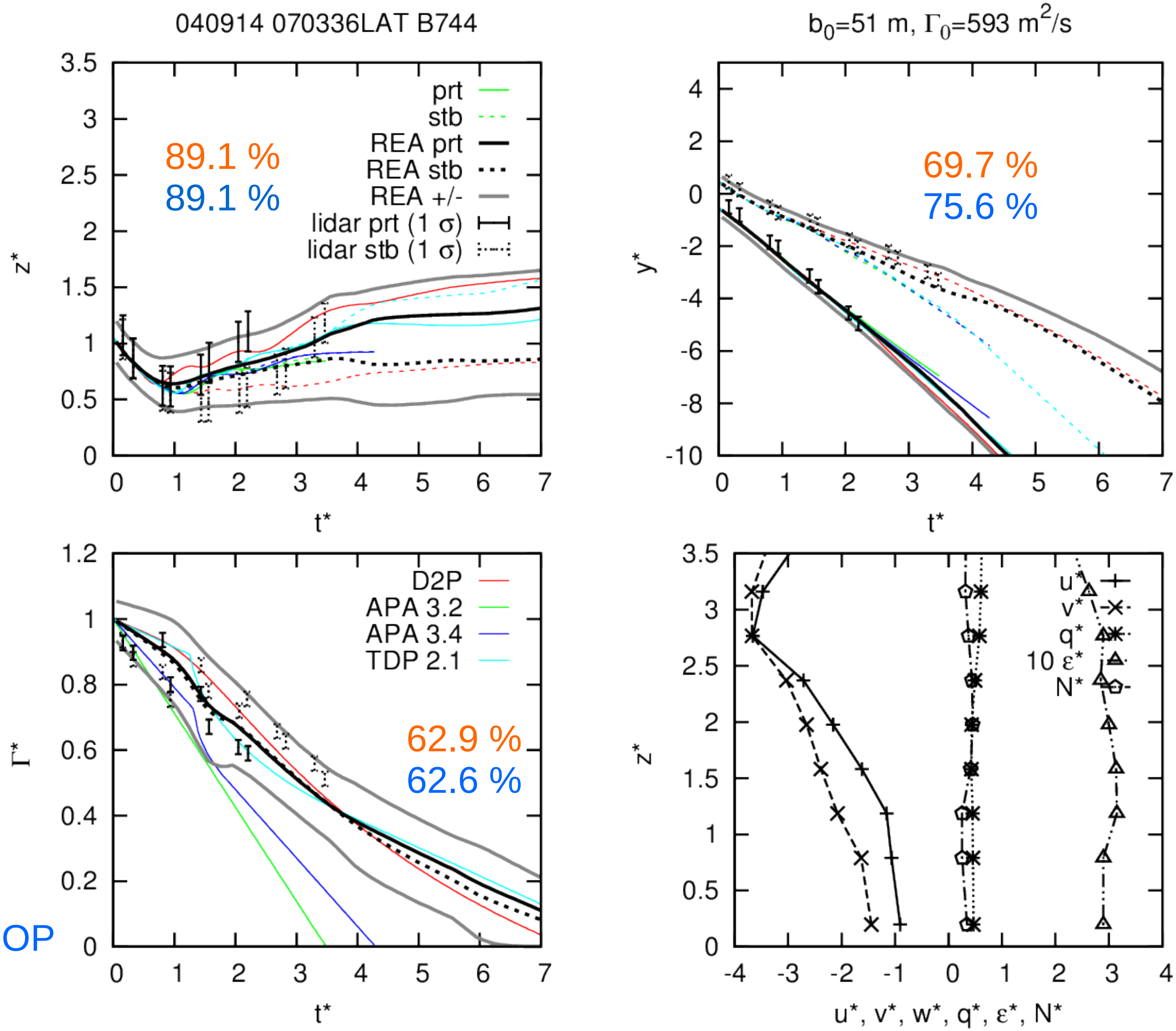
enhancement:

rmse $_{z^*,TDP}$=0.158
rmse $_{z^*,REA}$=0.148

rmse $_{\Gamma^*,D2P}$=0.085
rmse $_{\Gamma^*,REA}$=0.072

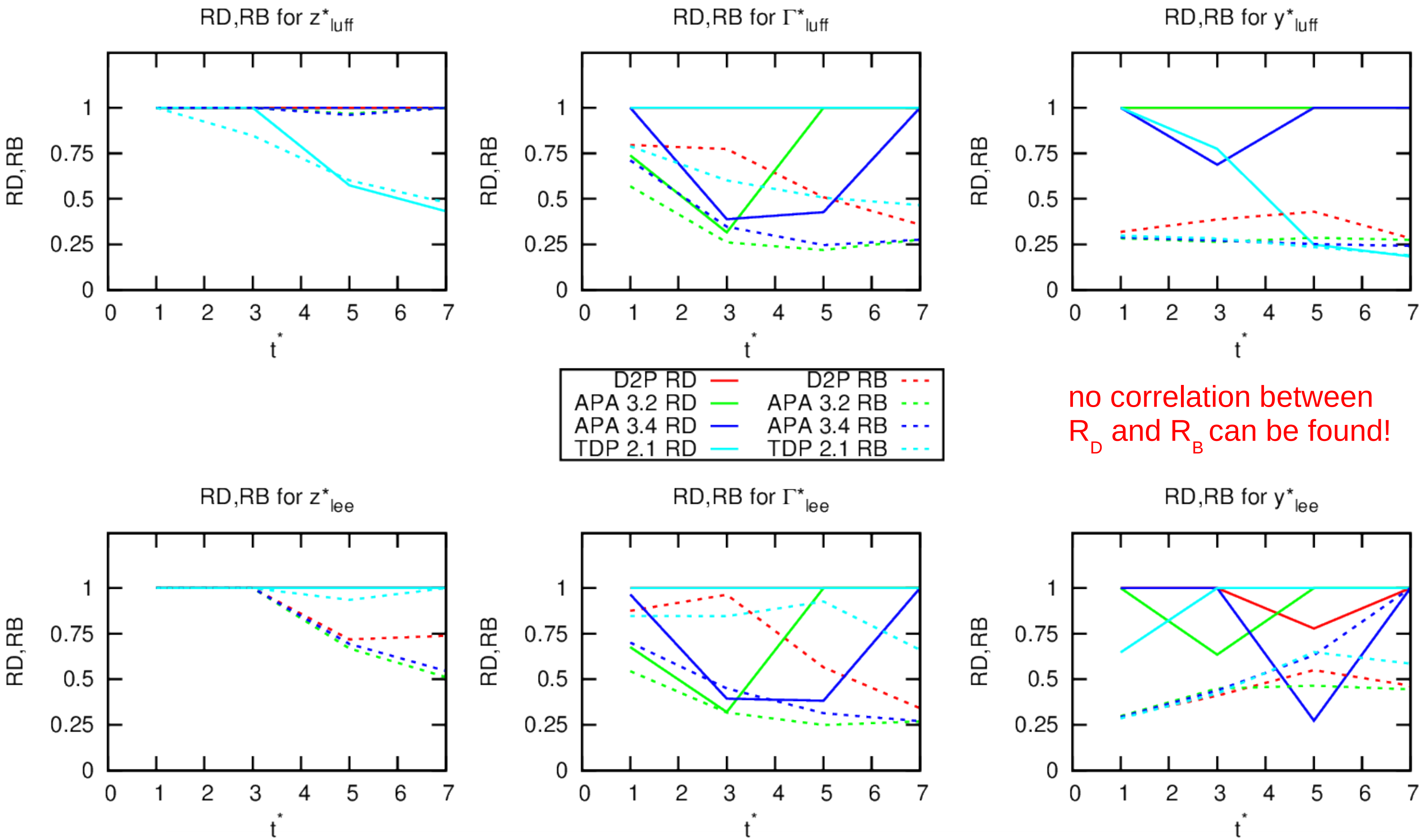
probability levels according to

- 99 testcases
- WakeFRA & WakeOP



Results

REA reliability factors (one single landing)



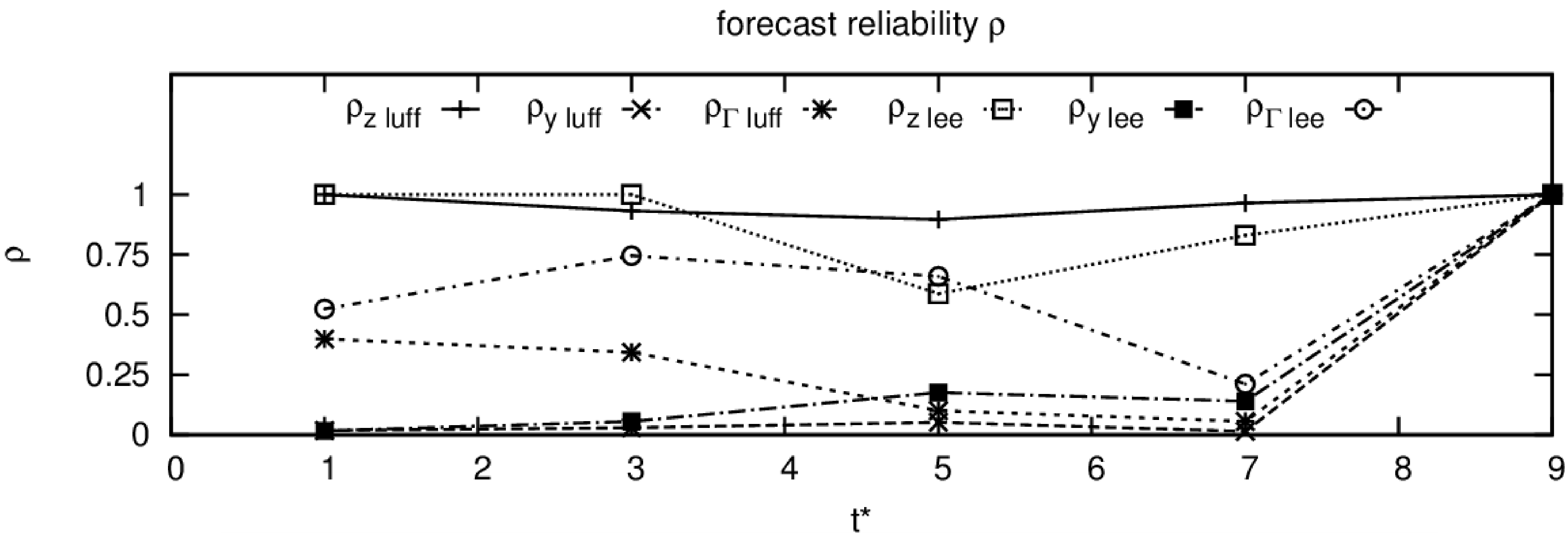
no correlation between R_D and R_B can be found!

Results

REA forecast reliability (one single landing)

$$\tilde{\rho} = \frac{\sum_i R_i^2}{\sum_i R_i}$$

- low reliability for y - forecast
- high reliability for z - forecast
- medium reliability for Γ - forecast



Results

REA scoring

- 99 randomly chosen cases
- skill factor s:

$$s_i = \frac{\sum_{p=1}^{n} rmse_{e,p}/rmse_{i,p}}{n} - 1$$

2[nd] best

best

median	rms Γ^*_{luff}	rms Γ^*_{lee}	rms y_{luff}	rms y^*_{lee}	rms z^*_{luff}	rms z^*_{lee}	s
REA	0.116	0.099	0.433	0.409	0.163	0.149	0.000
TDP 2.1	0.127	0.106	0.590	0.415	0.237	0.147	-0.122
APA 3.4	0.160	0.107	0.602	0.413	0.179	0.154	-0.127
APA 3.2	0.204	0.140	0.612	0.410	0.179	0.154	-0.190
D2P	0.122	0.120	0.406	0.408	0.140	0.166	-0.016

Results

REA scoring

- 99 randomly chosen cases
- skill factor s:

$$s_i = \frac{\sum_{p=1}^{n} rmse_{e,p}/rmse_{i,p}}{n} - 1$$

2nd best

best

median	rms Γ^*_{luff}	rms Γ^*_{lee}	rms y_{luff}	rms y^*_{lee}	rms z^*_{luff}	rms z^*_{lee}	s
REA	0.116	0.099	0.433	0.409	0.163	0.149	0.000
DEA	0.131	0.093	0.525	0.411	0.173	0.149	-0.048

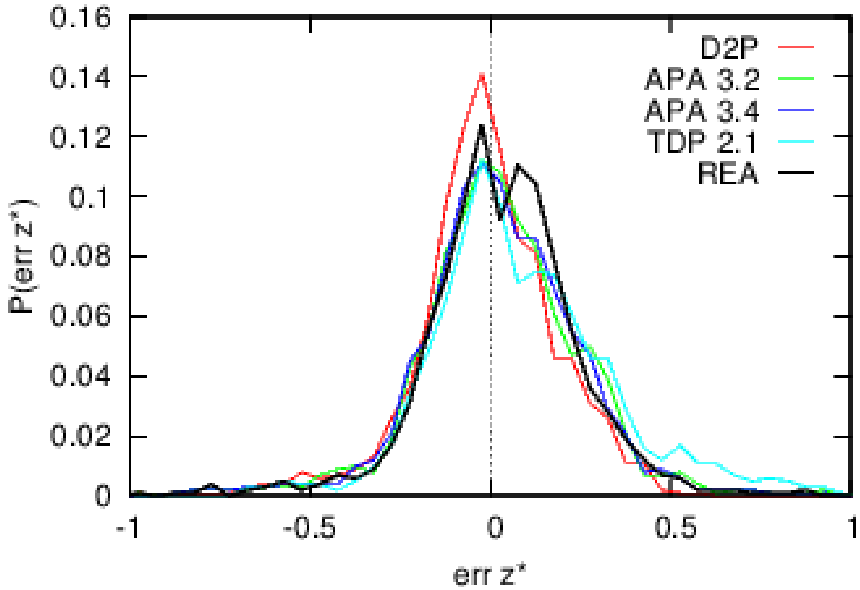
advanced MME approach outperforms Direct Ensemble Average (DEA)

PDD of models and ensemble

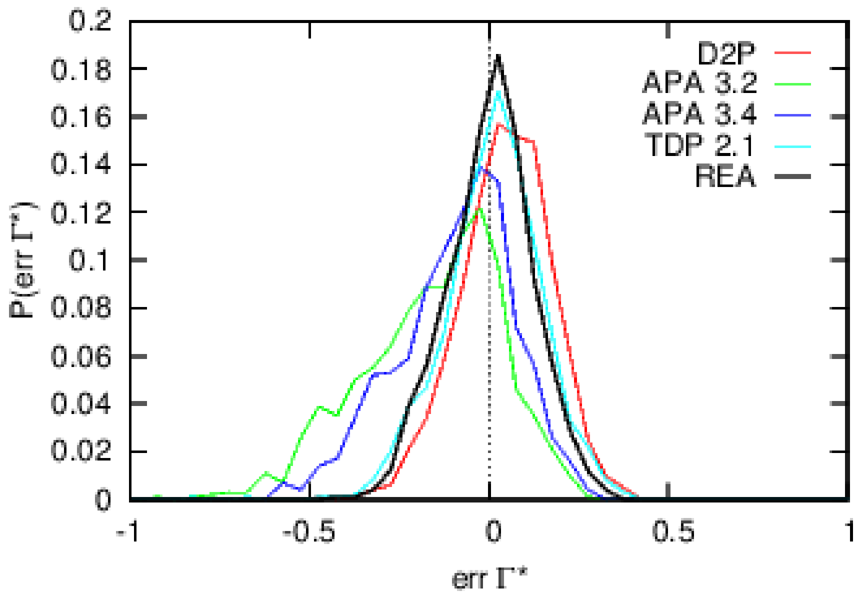
overconfident ensemble:
too narrow ensemble spread

well-dispersed ensemble:
coverage of full spectrum of possible solutions

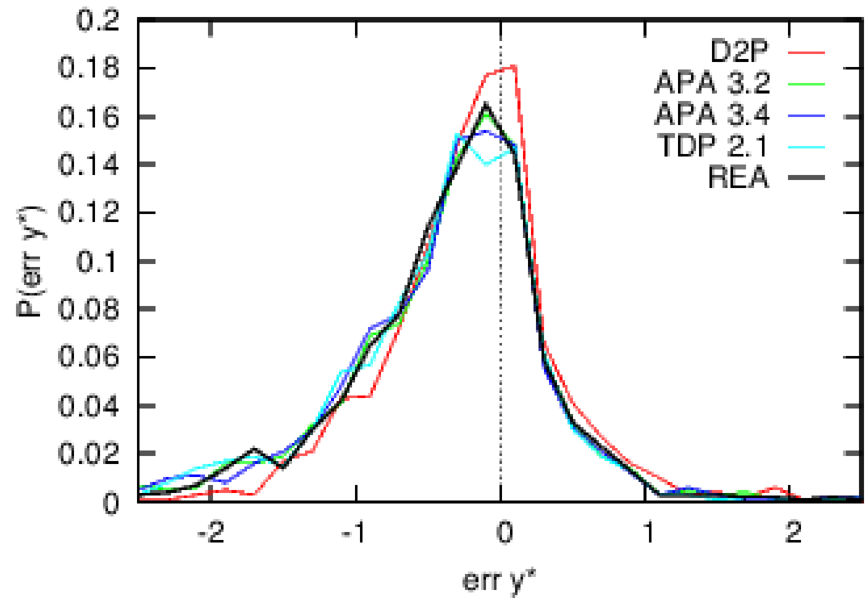
Weigel et al., 2008, Hagedorn et al.,2004



overconfident ensemble
→ small or no rmse improvement



well-dispersed model forecasts
→ rmse improvement



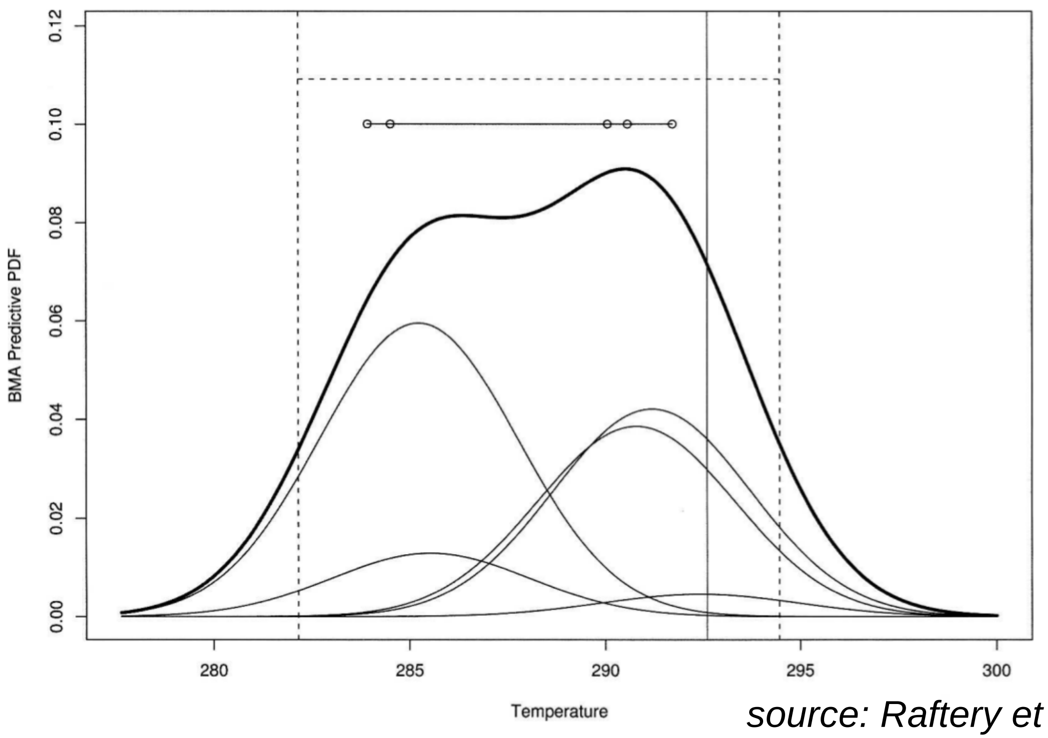
Conclusion

- ensemble **can improve quality** of wake vortex forecasts in average
- however **only 1.6 %** improvement compared to best model

  reason: ensemble is **overconfident** for z* and y*
- **but:** models might behave differently in particular ambient weather conditions and out-of-ground → investigation with pdds

Further Development

- How does **a good training data set** look like?
- Can the results be further improved by **distinguishing various ambient weather conditions**?
- How does the **Bayesian Model Averaging** (BMA) perform?

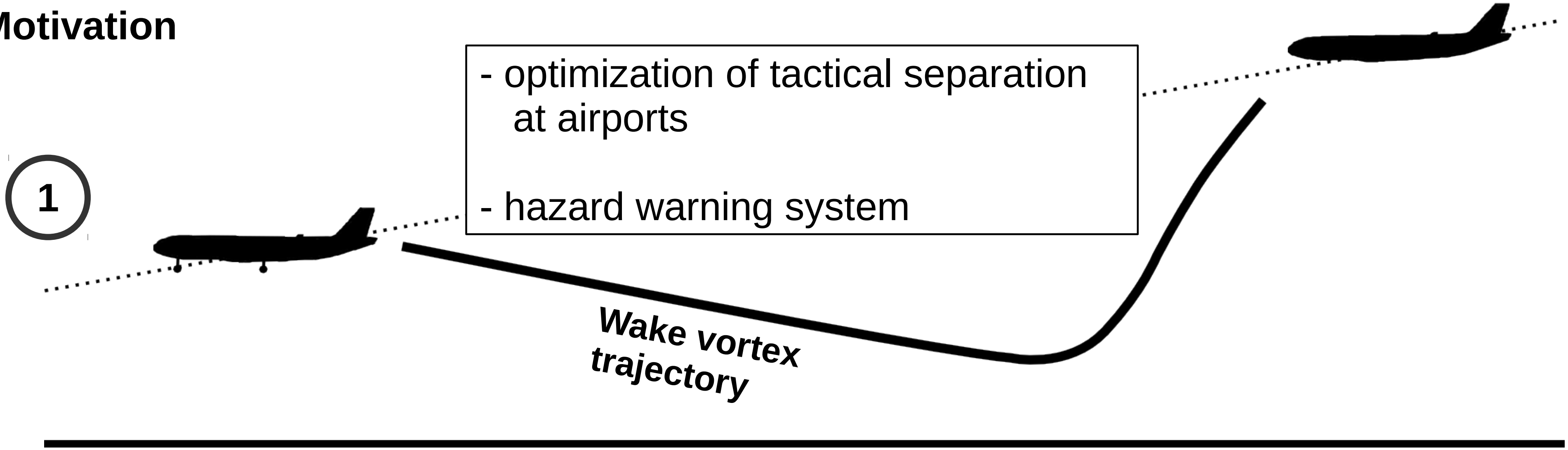


source: Raftery et al., 2005

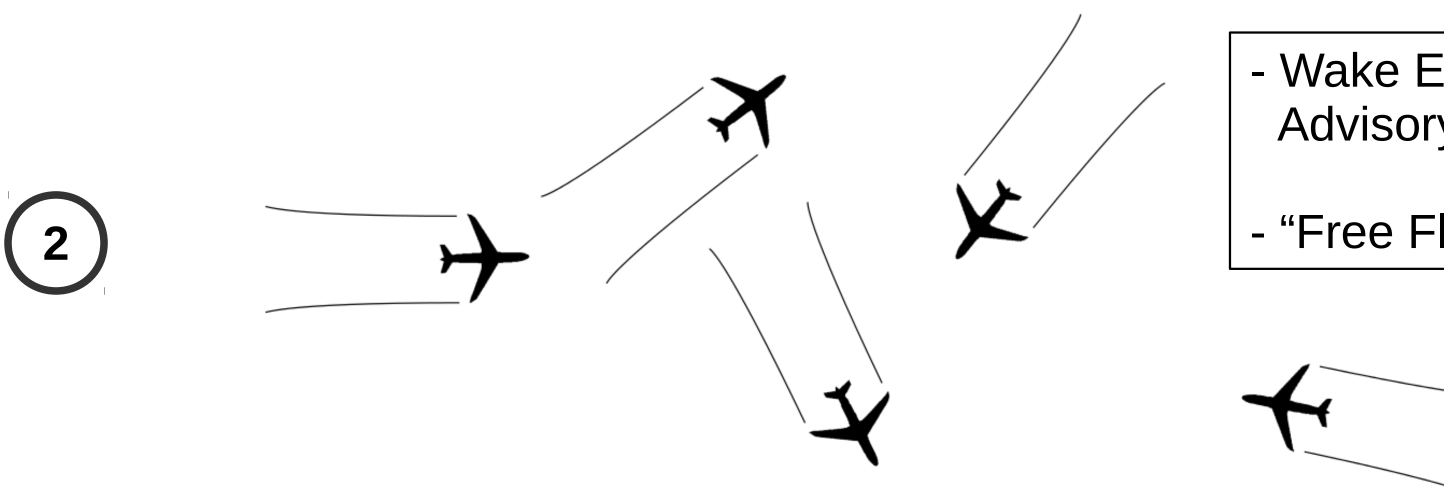
Backup

Wake Vortex Predictions

Motivation



- optimization of tactical separation at airports
- hazard warning system



- Wake Encounter Avoidance & Advisory System (WEAA)
- “Free Flight”

Ensemble Methoden
Bayesian Model Averaging

P(B) = Wahrscheinlichkeit des Eintretens von B

P(B|A) = Wahrscheinlichkeit für B, unter Vorraussetzung A

PDF = Probability Density Function (Wahrscheinlichkeitsdichtefunktion)

Law of total probability: $$P(B) = \sum_n P(B \cap A_n) = \sum_n P(A_n) P(B|A_n)$$

Beispiel:

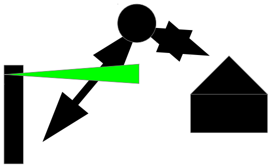
Wir befinden uns auf einem Schiff:
- wir wollen die Position B bestimmen
- 3 Crew-Mitglieder (A1,A2,A3) wissen wie es geht, haben aber unterschiedliche Methoden

according to Grimmett and Welsh., 1986

Ensemble Methoden

Bayesian Model Averaging

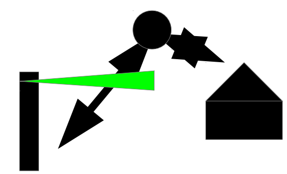
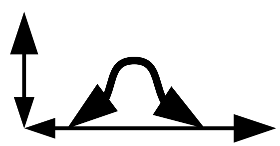
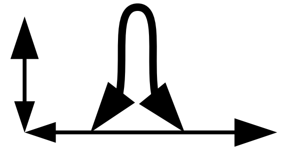
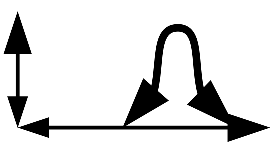
Law of total probability: $$P(B) = \sum_n P(B \cap A_n) = \sum_n P(A_n)P(B|A_n)$$

	A1	A2	A3
Methode			$\vec{s} = \sum_n \vec{v}_n * t_n$
individuelle Wahrscheinlichkeit, dass die Methode Erfolg hat: $P(B\|A_n)$	0.6	0.9	0.7
Wahrscheinlichkeit, dass wir A1, A2 or A3 fragen: $P(A_n)$	0.2	0.5	0.3

→ P(B)=0.78

Ensemble Methoden
Bayesian Model Averaging

Law of total probability: $$P(B) = \sum_n P(B \cap A_n) = \sum_n P(A_n) P(B|A_n)$$

Methode	A1 	A2 	A3 $\vec{s} = \sum_n \vec{v}_n * t_n$
PDF der Methode (Modell-Unsicherheiten): $P(B\|A_n)$			
Wahrscheinlichkeit, dass wir A1, A2 or A3 fragen: $P(A_n)$	0.1	0.6	0.3

Ensemble Methoden

Bayesian Model Averaging

Law of total probability: $$P(B) = \sum_n P(B \cap A_n) = \sum_n P(A_n) P(B|A_n)$$

angewandt auf Vorhersage-Modelle:

Annahme: es gibt immer ein bestes Ensemble-Glied

A_n = Modell n

B = vorherzusagende Größe

B^T = Trainings-Daten

$P(A_n)$ = Wahrscheinlichkeit, dass A_n das beste Modell ist

(***Gewichtungsfaktor***, basierend auf B^T)

$P(B|A_n)$ = ***PDF*** of A_n alone (Gaussian distribution, given that A_n is the best forecast)

≙ gewichtete Summe von Wahrscheinlichkeitsdichtefunktionen (PDFs)

according to Raftery et al., 2005

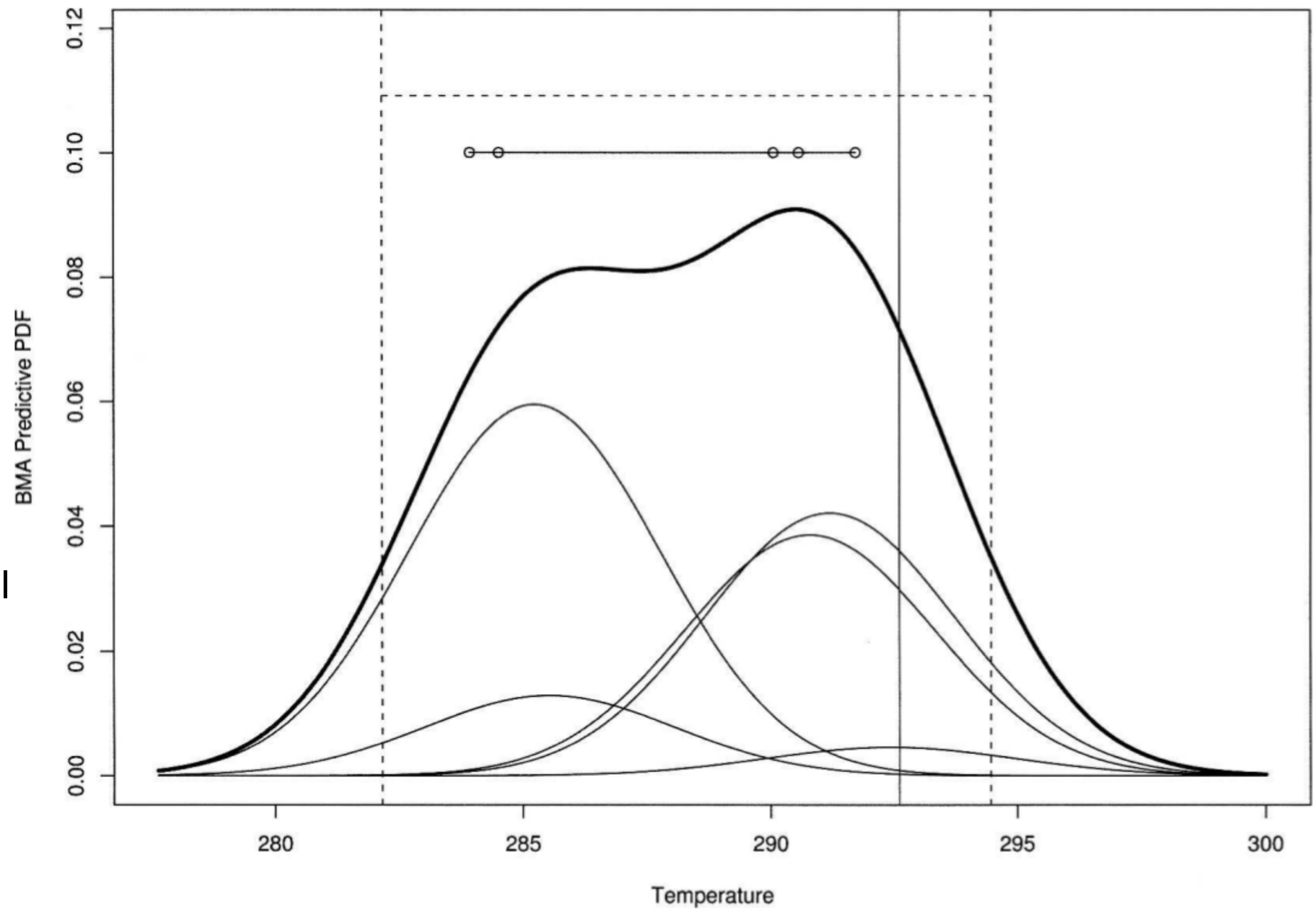
DLR

Ensemble Methoden

Bayesian Model Averaging

BMA applied on 48-h surface temperature forecast (bias corrected)

- ensemble forecast
- individual model PDF
- individual model forecast
- 90% interval
- verification

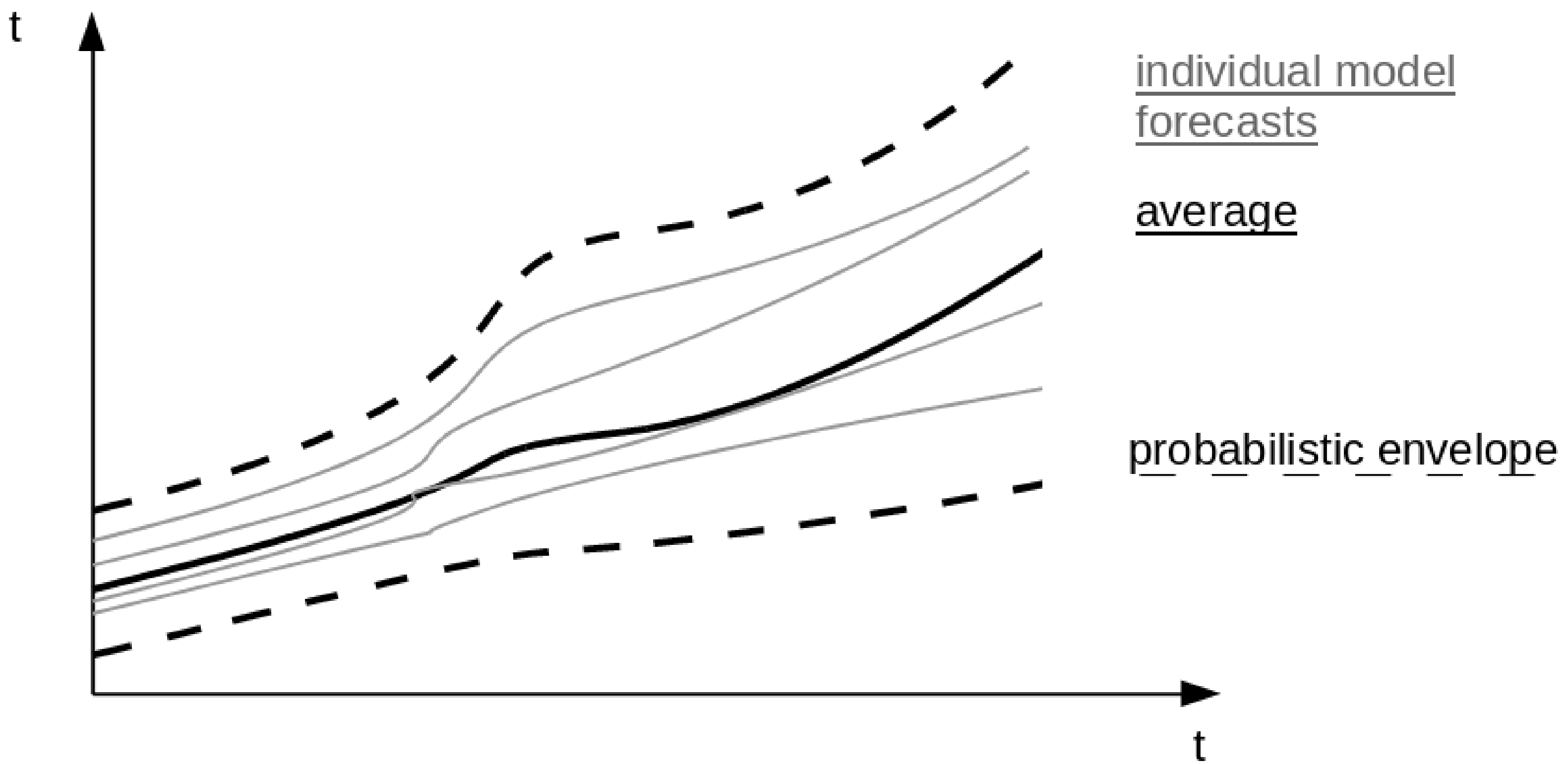


source: Raftery et al., 2005

Multi-Model Ensemble

benefit

- increase deterministic skill
- predict forecast skill
- provide probabilistic forecast



Multi-Model Ensemble